JOURNAL OF MULTIVARIATE ANALYSIS 18, 46–51 (1986)

The Multivariate Laplace–De Moivre Theorem

JERRY ALAN VEEH

Auburn University

Communicated by M. M. Rao

We find necessary and sufficient conditions for the convergence of (matrix) normed and centered k-variate binomial vectors to a full k-variate normal law. The norming matrices are also characterized. A converse to the univariate Laplace–De Moivre Theorem is obtained.

INTRODUCTION

Suppose we have a game some of whose possible outcomes consist of the events $E_1,..., E_k$. Let X be the k-dimensional random vector whose ith coordinate counts the number of occurrences of the event E_i in n independent plays of the game. We say that X has the k-variate binomial distribution based on n trials. Note that in the special case in which $E_1,..., E_k$ are mutually exclusive and exhaustive events X has the multinomial distribution. The multivariate binomial distribution has been studied in connection with discrete discriminant analysis where the components of X typically represent the number of affirmative responses obtained when n people complete questionaires consisting of k two-answer questions. (See [5].)

In this paper we study the asymptotic behavior of a sequence $\{X_n\}$ of k-variate binomial random vectors. We obtain an extension of the univariate Laplace–De Moivre theorem by finding necessary and sufficient conditions for the existence of a sequence of $k \times k$ matrices $\{A_n\}$ and k-dimensional vectors $\{a_n\}$ so that $A_n X_n + a_n$ converges in distribution to a full k-dimensional Gaussian law. The present work complements the results of Hudson, Tucker, and Veeh [3], which characterize some of the possible limit laws of such normed and centered multivariate binomial sequences.

To illustrate some of the consequences of matrix normalization and the requirement of a full limit we consider the

Received December 6, 1983; revised April 13, 1984.

AMS 1980 subject classifications: primary 60F05, 60E05.

Key words and phrases: Laplace–De Moivre theorem, binomial vectors, discriminant analysis

0047-259X/86 $3.00

EXAMPLE. Suppose that we have a game with two outcomes A and B with the property that in n independent plays of the game we have $P[A \cap B] = \frac{1}{2}$ and $P[A \backslash B] = P[B \backslash A] = 1/n$ for $n \geq 4$. Let X_n and Y_n, respectively, denote the number of occurrences of A and B in n plays of the game. Then

$$\begin{bmatrix} X_n \\ Y_n \end{bmatrix}$$

is bivariate binomial and one readily verifies that

$$n^{-1/2} \begin{bmatrix} X_n - \frac{1}{2}n \\ Y_n - \frac{1}{2}n \end{bmatrix} \to Z,$$

where Z is bivariate normal with mean vector 0 and covariance matrix

$$\frac{1}{4} \begin{bmatrix} 1 & 1 \\ 1 & 1 \end{bmatrix}.$$

Thus Z is *not* full. A short computation reveals that

$$\begin{bmatrix} 1 & -1 \\ 0 & n^{-1/2} \end{bmatrix} \begin{bmatrix} X_n \\ Y_n \end{bmatrix} - \begin{bmatrix} 0 \\ \frac{1}{2}n^{1/2} \end{bmatrix} \to \begin{bmatrix} U \\ V \end{bmatrix},$$

where U and V are independent random variables, U being the difference of two independent Poisson random variables and V being normal. Hence this normalization gives a full limit. Moreover, Billingsley's convergence of types theorem shows that *any full* limit law arising from matrix normalization of

$$\begin{bmatrix} X_n \\ Y_n \end{bmatrix}$$

must be a nonsingular affine linear transformation of

$$\begin{bmatrix} U \\ V \end{bmatrix}.$$

Thus a non-trivial Poisson component must be present in any full limit law. This type of behavior is explored more fully in [3]. In this paper we characterize those multivariate binomial systems which can be normalized to give full Gaussian limits.

Before proceeding we make a few formal definitions.

DEFINITION. Let $\mathscr{U}$ denote the vertices of the unit cube in R^k. A k-dimensional random vector X is said to be *Bernoulli* if $P[X \in \mathscr{U}] = 1$. A k-dimensional random vector Y is said to have the *k-dimensional binomial distribution based on n trials* if Y has the same distribution as the sum of n independent identically distributed k-dimensional Bernoulli random vectors.

We also recall that a random vector Z is *full* if there is no hyperplane H with the property that $P[Z \in H] = 1$. Observe that if EZ exists Z is full if and only if $0 < \mathrm{Var}\langle t, Z\rangle \leqslant +\infty$ for all vectors t with $\|t\| = 1$. Here $\langle \cdot, \cdot \rangle$ and $\|\cdot\|$ are the usual inner product and norm on R^k.

A final word on notation. If X_n converges to X in distribution we shall write $X_n \to X$. If X and Y have the same distribution we shall write $X = Y$. Finally since subsequencing arguments are used repeatedly here we shall *not* use double subscript notation but shall rely on the context to make the situation plain.

THE THEOREM

With the notation and definitions above we may now state the

THEOREM. *Let $\{X_n\}$ be a sequence of k-dimensional binomial random vectors. Let C_n denote the covariance matrix of X_n. The following are equivalent*:

(i) *There exist $k \times k$ matrices $\{A_n\}$ and k-dimensional vectors $\{a_n\}$ so that $A_n X_n + a_n \to Z$, where Z is a full k-dimensional Gaussian random vector.*

(ii) $\inf_{\|t\|=1} \langle t, C_n t\rangle \to +\infty$ *as* $n \to \infty$.

(iii) $\|C_n^{-1}\| \to 0$ *as* $n \to \infty$.

(iv) $C_n^{-1/2}(X_n - EX_n)$ *converges in distribution to the standard k-variate normal random vector (with zero mean and identity covariance matrix).*

Remark. In fact the theorem holds as long as each X_n is the sum of n independent identically distributed random vectors all of which take values in a fixed finite set.

Proof. We first show that (i) implies (ii). To do this we first symmetrize. Let us write X^s for the symmetrization of the random vector X. We have $(A_n X_n + a_n)^s = A_n X_n^s$. We also write $X_n = X_{n1} + \cdots + X_{nk_n}$, where

$X_{n1}, ..., X_{nk_n}$ are independent identically distributed k-variate Bernoulli random vectors. Thus

$$A_n X_n^s = \sum_{j=1}^{k_n} A_n X_{nj}^s \to Z^s \tag{1}$$

and Z^s is a full Gaussian random vector since Z is. By a result of R. Ranga Rao (see [6]), (1) implies that

$$\sup\{ |P[A_n X_n^s \in C] - P[Z^s \in C]| : C \in \mathscr{C}\} \to 0 \tag{2}$$

as $n \to \infty$, where $\mathscr{C}$ is the class of all measurable convex subsets of R^k. Since the covariance matrix of X_n^s is $2C_n$ it suffices to show that (1) and (2) imply (ii).

Our first claim is that $k_n \to \infty$ as $n \to \infty$. If this were not the case upon passing to a subsequence we would have

(a) $k_n \leqslant M$, some $0 < M < \infty$, and

(b) $A_n X_n^s \to Z^s$.

But X_n^s has at most $3^{k\,k_n} \leqslant 3^{Mk}$ possible values, so for each n, $A_n X_n^s$ has an atom of mass at least 3^{-Mk}, contradicting (2). Thus $k_n \to \infty$ as $n \to \infty$.

Suppose now that condition (ii) fails. Again passing to a subsequence we can find unit vectors t, t_1, t_2,... and a number R so that

(a) $\langle t_n, C_n t_n \rangle \leqslant R$ for all n,

(b) $t_n \to t$ as $n \to \infty$, and

(c) $A_n X_n^s \to Z^s$.

Let $\omega \in \mathscr{U} - \mathscr{U}$ be a possible value of X_{n1}^s such that $\langle t, \omega \rangle \neq 0$. Then for large n

$$\begin{aligned} \langle t, \omega \rangle^2 P[X_{n1}^s = \omega] &\leqslant 2\langle t_n, \omega \rangle^2 P[X_{n1}^s = \omega] \\ &\leqslant 2E[\langle t_n, X_{n1}^s \rangle^2] \\ &\leqslant (2/k_n) E[\langle t_n, X_n^s \rangle^2] \\ &= (4/k_n) \langle t_n, C_n t_n \rangle \\ &\leqslant 4R/k_n. \end{aligned}$$

Thus for large n, $P[\langle t, X_{n1}^s \rangle = 0] \geqslant 1 - Dk_n^{-1}$ for some constant D independent of n. Hence

$$\varliminf_{n \to \infty} P[\langle t, X_n^s \rangle = 0] \geqslant \varliminf_{n \to \infty} (1 - Dk_n^{-1})^{k_n} = e^{-D}$$

so for large n, X_n^s (and hence $A_n X_n^s$) assigns mass at least $e^{-D}/2$ to a $(k-1)$-dimensional subspace. This contradicts (2) and so (ii) holds.

Since C_n is non-negative definite it is clear that (ii) implies (iii).

We now show that (iii) implies (iv). Since $\| C_n^{-1} \| \to 0$ we have $\| C_n^{-1/2} \| \to 0$. Thus the system $\{C_n^{-1/2} X_{nj};\ 1 \leqslant j \leqslant k_n,\ n = 1, 2, ...\}$ is infinitesimal. Let t, $\|t\| = 1$, be arbitrary. We have, for any $\varepsilon > 0$, $k_n P[\,|\langle t, C_n^{-1/2} X_{n1} \rangle| > \varepsilon] = 0$ for large n since X_{n1} takes values only in $\mathscr{U}$ and $\| C_n^{-1/2} \| \to 0$. In fact for large n, $\{\,|\langle t, C_n^{-1/2} X_{n1} \rangle| > \varepsilon\} = \varnothing$. Thus

$$\lim_{\varepsilon \downarrow 0} \overline{\lim_{n}}\, k_n \operatorname{Var}\langle t, C_n^{-1/2} X_{n1} \rangle\, 1_{\{|\langle t, C_n^{-1/2} X_{n1} \rangle| < \varepsilon\}}$$

$$= \lim_{\varepsilon \downarrow 0} \overline{\lim_{n}}\, k_n \operatorname{Var}\langle t, C_n^{-1/2} X_{n1} \rangle = 1.$$

Thus by the central limit theorem (see [4]), $\langle t, C^{-1/2}(X_n - EX_n) \rangle$ converges to the standard normal law for each t. Hence $C_n^{-1/2}(X_n - EX_n)$ converges to the standard k-variate normal law. Since (iv) clearly implies (i) the theorem is proved.

Again using the results of Billingsley [1] we obtain the

COROLLARY. *Let* (i) *of the theorem hold and let Σ be the positive definite covariance matrix of Z. Then $A_n = I_n \Sigma^{1/2} O_n C_n^{-1/2}$, where $\{O_n\}$ is a sequence of orthogonal matrices and $\{I_n\}$ is a sequence of non-singular matrices such that $I_n \to I$, the $k \times k$ identity matrix.*

We note in conclusion that the theorem provides a converse to the usual univariate Laplace–De Moivre theorem (see, e.g. Feller [2]). Thus in the univariate case $(\operatorname{Var} X_n)^{-1/2}(X_n - EX_n)$ converges to the standard normal distribution if and only if $\operatorname{Var} X_n \to \infty$.

ACKNOWLEDGMENT

The author thanks the referee for several suggestions which improved the quality of this paper.

REFERENCES

[1] BILLINGSLEY, P. (1966). Convergence of types in k-space. *Z. Wahrsch. Verw. Gebiete* **5** 175–179.

[2] FELLER, W. (1968). *An Introduction to Probability and Its Applications*, Vol. 1, 3rd ed., p. 186, Wiley, New York.

[3] Hudson, W., Tucker, H., and Veeh, J. Limit distributions of multivariate binomial systems. *J. Multivariate Anal.*, to appear.

[4] Gnedenko, B., and Kolmogorov, A. (1968) *Limit Distributions for Sums of Independent Random Variables*, p. 116, Addison–Wesley, Reading, Mass.

[5] Goldstein, M., and Dillon, W. R. (1978). *Discrete Discriminant Analysis.* Wiley, New York.

[6] Rao, R. R. (1962). Relations between weak and uniform convergence of measures with applications. *Ann. Math. Statist.* **33** 659–680.